

Table 12: IRES Sequences

SEQ ID NO:___ A 519 base pair IRES obtainable from encephelomyocarditis virus (EMCV).

1 GACGTCGACTAAATCCGGTTATTTTCCACCATATTGCCGTCTTTTGGCAA
 SalI
 51 TGTGAGGGCCCCGAAACCTGGCCCTGTCTTCTTGACGAGCATTCTAGGG
 101 GTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAG
 151 GAAGCAGTTCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGAC
 201 CCTTTGCGAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCC
 251 AAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCAGTGC
 301 CACGTTGTGAGTTGGATAGTTGTGAAAGAGTCAAATGGCTCTCCTCAAG
 351 CGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTTGTATGG
 401 GATCTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGG
 451 TTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTTCCTTTGA
 SalI
 501 AAAACACGATGTCGACGTC

SEQ ID NO:___ An IRES obtainable from vascular endothelial growth factor (VEGF).

1 ACGTAGTTCGACAGCGCAGAGGCTTGGGGCAGCCGAGCGGCAGCCAGGCCC
 SalI
 51 CGGCCCGGGCCTCGGTTCCAGAAGGGAGAGGAGCCCGCAAGGCGCGCAA
 101 GAGAGCGGGCTGCCTCGCAGTCCGAGCCGAGAGGGAGCGCGAGCCGCGC
 151 CGGCCCGGACGGCCTCCGAAACCATGGTTCGACACGTA
 SalI

SEQ ID NO:___ A 5'UTR region of HCV.

1
 61 GCCAGCCCCCTGATGGGGGCGACACTCCGCCATGAATCACTCCCCTGTGAGGAACTACTG
 121 TCTTCACGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGAC
 181 CCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG
 241 GACGACCGGGTCTTTCTTGGATTAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCC
 301 GCAAGACTGCTAGCCGAGTAGTGTGGGTGCGGAAAGGCCTTGTGGTACTGCCTGATAGG
 GTGCTTGCAGGTGCCCCGGGAGGTCTCGTAGACCGTGCACC (341)

SEQ ID NO: __ A 5'UTR region of BiP SEQ ID NO:4

1
CCCCGGGGTCACTCCTGCTGGACCTACTCCGACCCCCTAGGCCGGGAGTGAAGGCGGGACT
61
5 TGTGCGGTTACCAGCGGAAATGCCTCGGGGTCAGAAGTCGCAGGAGAGATAGACAGCTGC
121 TGAACCAATGGGACCAGCGGATGGGGCGGATGTTATCTACCATTGGTGAACGTTAGAAAC
181 GAATAGCAGCCAATGAATCAGCTGGGGGGGCGGAGCAGTGACGTTTATTGCGGAGGGGGC
241 CGCTTCGAATCGGCGGCGGCCAGCTTGGTGGCCTGGGCCAATGAACGGCCTCCAACGAGC
301 AGGGCCTTCACCAATCGGCGGCCTCCACGACGGGGCTGGGGGAGGGTATATAAGCCGAGT
10 361 AGGCGACGGTGAGGTGCAGCGCCGCCAAGACAGCACAGACAGATTGACCTATTGGGGTGT
421 TTCGCGAGTGTGAGAGGGAAGCGCCGCGGCCTGTATTTCTAGACCTGCCCTTCGCTTGGT
481 TCGTGGCGCCTTGTGACCCCCGGGCCCCCTGCCGCTGCAAGTCGAAATTGCGCTGTGCTCC
541 TGTGCTACGGCCTGTGGCTGGACTGCCTGCTGCTGCCCAACTGGCTGGCAAGATG (595)

SEQ ID NO: __ A 5'UTR of PDGF SEQ ID NO:5

1
GTTTGACCTCTCCCTGCCCCGGGTGCTCGAGCTGCCGTTGCAAAGCCAAC'TTTGGAAAAA
61
GTTTTTTGGGGGAGACTTGGGCCTTGAGGTGCCAGCTCCGCGCTTCCGATTTTGGGGG
20 121 CTTTCCAGAAAATGTTGCAAAAAAGCTAAGCCGGCGGGCAGAGGAAAACGCCCTGTAGCCG
181 GCGAGTGAAGACGAACCATCGACTGCCGTGTTCTTTTCTCTTGGAGGTTGGAGTCCCC
241 TGGGCGCCCCCACACCCCTAGACGCCTCGGCTGGTTCGCGACGCAGCCCCCGGCCGTGG
301 ATGCTGCACTCGGGCTCGGGATCCGCCAGGTAGCCGGCCTCGGACCCAGGTCTGCGCC
361 CAGGTCCTCCCCCTGCCCCCAGCGACGGAGCCGGGGCCGGGGGCGGCGGCGCGGGGGCA
25 421 TGCGGGTGAGCCGCGGCTGCAGAGGCCTGAGCGCCTGATCGCCGCGGACCTGAGCCGAGC
481 CCACCCCCCTCCCCAGCCCCCACCCTGGCCGCGGGGGCGGCGCGCTCGATCTACGCGTC
541 CGGGGCCCCGCGGGGCCGGGCCCCGGAGTCGGCATG (575)

Table 13: Literature References For IRES

IRES Host	Example	Reference
Picornavirus	HAV	Glass et al., 1993. Virology 193:842-852
	EMCV	Jang & Wimmer, 1990. Gene Dev 4:1560-1572
	Poliovirus	Borman et al., 1994. EMBO J 13:3149-3157
HCV and pestivirus	HCV	Tsukiyama-Kohara et al., 1992. J Virol 66:1476-1483
	BVDV	Frolov I et al., 1998. RNA. 4:1418-1435
Leishmania virus	LRV-1	Maga et al., 1995. Mol Cell Biol 15:4884-4889
Retroviruses	MoMLV VL30 (Harvey murine sarcoma virus)	Torrent et al., 1996. Hum Gene Ther 7:603-612
	REV	Lopez-Lastra et al., 1997. Hum Gene Ther 8:1855-1865
Eukaryotic mRNA	BiP	Macejak & Sarnow, 1991. Nature 353:90-94
	antennapedia mRNA	Oh et al., 1992. Gene & Dev 6:1643-1653
	FGF-2	Vagner et al., 1995. Mol Cell Biol 15:35-44
	PDGF-B	Bernstein et al., 1997. J Biol Chem 272:9356-9362
	IGFII	Teerink et al., 1995. Biochim Biophys Acta 1264:403-408
	eIF4G	Gan & Rhoads, 1996. J Biol Chem 271:623-626
	VEGF	Stein et al., 1998. Mol Cell Biol 18:3112-3119; Huez et al., 1998. Mol Cell Biol 18:6178-6190

Table 14: TRE Sequences

Nucleotide sequence of a human uroplakin II 5' flanking region. Position +1 (the translational start site) is denoted with an asterisk. SEQ ID NO: __ (number 1 of

SEQ ID NO: __ corresponds to position -2239 with respect to the translational start site).

5	TCGATAGGTA 1	CCCACTATAG	GGCACGCGTG	GTCGACGGCC	CGGGCTGGTC 50
	TGGCAACTTC 51	AAGTGTGGGC	CTTTCAGACC	GGCATCATCA	GTGTTACGGG 100
10	GAAGTCACTA 101	GGAATGCAGA	ATTGATTGAG	CACGGTGGCT	CACACCTGTA 150
	ATCCCAACAC 151	TCTGGGAGGC	CAAGGCAGGT	GGATCACTTG	TGGTCAGGAG 200
15	TTTGAGACCA 201	GCCTGGCCAA	CATGGTGAAA	CCTCATCTCT	ACTAAAAATA 250
	CAAAAATTAG 251	CTGGGAATGG	TGGCACATGC	CTATAATCCC	AGTTACTCAG 300
20	GAGGCTGAGG 301	CAGGAGAATC	ATTTGAACCT	GGGAGGCAGA	GGTTGCAGTG 350
	AGCCGAGATC 351	ACGCCACTGC	ACTCCAGCCT	GGGTGACACA	GCGAGACTCT 400
25	GTCTCAAAAA 401	AAAAAAAATG	CAGAATTTCA	GGCTTCACCC	CAGACCCACT 450
	GCATGACTGC 451	ATGAGAAGCT	GCATCTTAAC	AAGATCCCTG	GTAATTCATA 500
30	CGCATATTAA 501	ATTTGGAGAT	GCACTGGCGT	AAGACCCTCC	TACTCTCTGC 550
	TTAGGCCCAT 551	GAGTTCTTCC	TTTACTGTCA	TTCTCCACTC	ACCCCAAACCT 600
40	TTGAGCCTAC 601	CCTTCCCACC	TTGGCGGTAA	GGACACAACC	TCCCTCACAT 650
	TCCCTACCAGG 651	ACCCTAAGCT	TCCCTGGGAC	TGAGGAAGAT	AGAATAGTTC 700
45	GTGGAGCAAA 701	CAGATATACA	GCAACAGTCT	CTGTACAGCT	CTCAGGCTTC 750
	TGGAAGTTCT 751	ACAGCCTCTC	CCGACAAAGT	ATTCCACTTT	CCACAAGTAA 800
50	CTCTATGTGT 801	CTGAGTCTCA	GTTTCCACTT	TTCTCTCTCT	CTCTCTCTCT 850

	CACCAGAATG	TTTCTCTCCC	ATTCTTACCC	ACTCAAGGCC	CTTTCAGTAG	1800
	1751					
5	CCCCTTGGAG	TATTCTCTTC	CTACATATCA	GGGCAACTTC	CAAACCTCATC	1850
	1801					
	ACCCTTCTGA	GGGGTGGGGG	AAAGACCCCC	ACCACATCGG	GGGAGCAGTC	1900
	1851					
10	CTCCAAGGAC	TGGCCAGTCT	CCAGATGCCC	GTGCACACAG	GAACACTGCC	1950
	1901					
	TTATGCACGG	GAGTCCCAGA	AGAAGGGGTG	ATTTCTTTCC	CCACCTTAGT	2000
	1951					
	TACACCATCA	AGACCCAGCC	AGGGCATCCC	CCCTCCTGGC	CTGAGGGCCA	2050
	2001					
20	GCTCCCCATC	CTGAAAACC	TGTCTGCTCT	CCCCACCCCT	TTGAGGCTAT	2100
	2051					
	AGGGCCCAAG	GGGCAGGTTG	GA CTGGATT C	CCCTCCAGCC	CCTCCCGCCC	2150
	2101					
25	CCAGGACAAA	ATCAGCCACC	CCAGGGGCAG	GGCCTCACTT	GCCTCAGGAA	2200
	2151					
	CCCCAGCCTG	CCAGCACCTA	TTCCACCTCC	CAGCCCAGCA		2239
	2201					
30						

Nucleotide sequence of a mouse uroplakin II 5' flanking region. The translational start site is denoted with an asterisk. SEQ ID NO:___ (number 1 of SEQ ID NO:6___corresponds to position -3592 with respect to the translational start site).

	CTCGAGGATCTCGGCCCTCTTTCTGCATCCTTGTCCTAAATCATTTTCAT	50
	1	
40	ATCTTGCTAGACCTCAGTTTGAGAGAAACGAACCTTCTCATTTTCAAGTT	100
	51	
	GAAAAAAAAAAGAGGTTCAAAGTGGCTCACTCAAAGTTACAAGCCAACAC	150
	101	
45	TCACCACTACGAGTACAATGGCCACCATTAGTGCTGGCATGCCCCAGGAG	200
	151	
50	ACAGGCATGCATATTATTCTAGATGACTGGGAGGCAGAGGGGTGGCCTAG	250
	201	

TGAGGTCAGACTGTGGACAGATCAGGCAGATGTGGGTTCTGATCCCAATT
251 300

5 CCTCAGGCCGCAGAACTACTGTGGTTCAAGAAGGGGACAAAAGGACTGCA
301 350

GTCCGGAACAGGAGGTCCATTTGAGAGCTGACTGAGCAGAAGAGGAAAGT
351 400

10 GAAGAACTTCTGGGGCAAGAGCTTACCCTACTTTACAGCTTTGTTGTCTT
401 450

15 CTTTACTCCAGGGGCGTCCCTGGTACTCAGTAAATGTCTGTTGGCTTGAG
451 500

GAACATATGTGTAAGGAGGAAGGAGAGGGAAGTTGAGGGAGTTAAGACTC
501 550

20 AAGAATCAATCAAGGAGAGGACAGCAGAGAAGACAGGGTTTGGGAGAGAG
551 600

ACTCCAGACATTGGCCCTGGTTCCTTCTTGCCACTGTGAAACCCTCCA
601 650

25 GAGGAACTGAGTGCTGTGGCTTTAAATGATCTCAGCACTGTCAGTGAAGC
651 700

GCTCTGCTCAAAGAGTTATCCTCTTGCTCCTGTGCCGGGGCCTCCCCCTC
701 750

30 CTCTCAGCTCCCAAACCCTTCTCAGCCACTGTGATGGCATAATTAGATGC
751 800

35 GAGAGCTCAGACCGTCAGGTCTGCTCCAGGAACCACCCATTTTCCCCAAC
801 850

CCCAGAGAAAGGTCCTAGTGGAAGAGTGGGGGCACTGAAGGGCTGATGG
851 900

40 GGTTCTGTCCTTTCCCCCATGCTGGGTGGACTTAAAGTCTGCGATGTGTG
900 950

TAGGGGGTAGAAGACAACAGAACCTGGGGGCTCCGGCTGGGAGCAGGAGG
951 1000

45 AACTCTCACCAGACGATCTCCAAATTTACTGTGCAATGGACGATCAGGAA
1001 1050

50 ACTGGTTCAGATGTAGCTTCTGATACAGTGGGTCTGAGGTAAAACCCGAA
1051 1100

ACTTAATTTCTTTCAAAAATTTAAAGTTGCATTTATTATTTTATATGTGT
1101 1150

5 GCCCATATGTGTGCCACAGTGTCTATGTGGAGGTCAGAGGGCAAGTTGTG
1151 1200

GGCATTGGCTCTCTCCTTTTCATAATGTGGCTTCTGGGGACCAAAATGTCA
1201 1250

10 GGCATGGTGGCAAGAGCTTTTACCTGTTGAGCCATCTCATGGTTTCGTAA
1251 1300

15 AACTTCCTATGACGCTTACAGGTAACGCAGAGACACAGACTCACATTTGG
1301 1350

AGTTAGCAGATGCTGTATTGGTGTAAACACTCATACACAGACACACACAC
1351 1400

20 ATACTCATACACACACACACACTTATCACATGCACACACATACTCGTA
1401 1450

TACACACAGACACACACACATGCACTCTCACATTTCACATATTCATACACA
1451 1500

25 TCCACACACACACTCATCCACACACACAGACACACATACTCATCCACACA
1501 1550

30 CACACACACACATACTCATACACACACAGACACACATACTCATACACA
1551 1600

CACACAGACACACACATATAATCATACATACACAGACACACTCATACATG
1601 1650

35 TGCACACACACACTCATCCACACACACACACTCATACACACACACTCA
1651 1700

TACACACACACACTCATACACACACACACGAGGTTTTTCTCAGGCTGCCT
1701 1750

40 TTGGGTGGAGACTGGAAGTGAATTTCTGTTTTTCAGCTCCTTGGCTTTTTG
1751 1800

45 TCCCTTTAGATGAGATCTCCTCCTCACTTTACACACAGAAAGATCACACA
1801 1850

CGAGGGAGAACTGGCGGTGCGGAAGAGGGCTACACGGTAGGGTGTGAGGG
1851 1900

50 TCAGGAGATCTTCCTGGCAAGTCTCAAACCTCCACATAGCACAGTGTTTA

	1901	1950
	CGTGAGGATTTAGGAGGAATCAGGAAGAGGATTGGTTTACTGCAGAGCAG	
	1951	2000
5	ACCATATAGGTCCACTCCTAAGCCCCATTTGAAATTAGAAGTGAGACAGT	
	2001	2050
	GTGGGATAAAAAGAGCAGATCTCTGGTCACATTTTTAAAGGGATATGAGG	
10	2051	3000
	GTCCTGTGCCTTTAAGCCTTCCCATCTCCCTCCAATCCCCCCTCACCTTC	
	2101	2150
	CCCACCCTAACCCCTCCCCAGGTTTCTGGAGGAGCAGAGTTGCGTCTTCTC	
15	2151	2200
	CCTGCCCTGCCGAGCTGCTCACTGGCTGCTCTAGAGGCTGTGCTTTGCGG	
	2201	2250
20	TCTCCATGGAAACCATTAGTTGCTAAGCAACTGGAGCATCATCTGTGCTG	
	2251	2300
	AGCTCAGGTCCTATCGAGTTCACCTAGCTGAGACACCCACGCCCCCTGCAG	
25	2301	2350
	CCACTTTGCAGTGACAAGCCTGAGTCTCAGGTTCTGCATCTATAAAAACG	
	2351	2400
30	AGTAGCCTTTTCAGGAGGGCATGCAGAGCCCCCTGGCCAGCGTCTAGAGGA	
	2401	2450
	GAGGTGACTGAGTGGGGCCATGTCACTCGTCCATGGCTGGAGAACCTCCA	
	2451	2500
35	TCAGTCTCCCAGTTAGCCTGGGGCAGGAGAGAACCAGAGGAGCTGTGGCT	
	2501	2550
	GCTGATTGGATGATTTACGTACCCAATCTGTTGTCCCAGGCATCGAACCC	
40	2551	2600
	CAGAGCGACCTGCACACATGCCACCGCTGCCCCGCCCTCCACCTCCTCTG	
	2601	2650
45	CTCCTGGTTACAGGATTGTTTTGTCTTGAAGGGTTTTGTTGTTGCTACTT	
	2651	2700
	TTTGCTTTGTTTTTTCTTTTTTAACATAAGGTTTCTCTGTGTAGCCCTAG	
50	2701	2750

CTGTCCTGGAAC TCACTCTGTAGACCAGGCTGGCCTCAAAC TCAAGAAATC
2751 2800

5 CACCTTCCTCCCAAGTGCTGGGATTAAAGGCATTTCGCACCATCGCCCAGC
2801 2850

CCCCGGTCTTGTTTCCTAAGGTTTTCTTGCTTTACTCGCTACCCGTTGCA
2851 2900

10 CAACCGCTTGCTGTCCAAGTCTGTTTGTATCTACTCCACCGCCCACTAGC
2901 2950

CTTGCTGGACTGGACCTACGTTTACCTGGAAGCCTTCACTAACTTCCCTT
2951 3000

15 GTCTCCACCTTCTGGAGAAATCTGAAGGCTCACACTGATACCCTCCGCTT
3001 3050

20 CTCCCAGAGTCGCAGTTTCTTAGGCCTCAGTTAAATACCAGAATTGGATC
3051 3100

TCAGGCTCTGCTATCCCCACCCTACCTAACCAACCCCTCCTCTCCCATC
3101 3150

25 CTTACTAGCCAAAGCCCTTTCAACCCTTGGGGCTTTTCCTACACCTACAC
3151 3200

ACCAGGGCAATTTTAGAACTCATGGCTCTCCTAGAAAACGCCTACCTCCT
3201 3250

30 TGGAGACTGACCCTCTACAGTCCAGGAGGCAGACACTCAGACAGAGGAAC
3251 3300

35 TCTGTCCTTCAGTCGCGGGAGTTCCAGAAAGAGCCATACTCCCCTGCAGA
3301 3350

GCTAACTAAGCTGCCAGGACCCAGCCAGAGCATCCCCCTTTAGCCGAGGG
3351 3400

40 CCAGCTCCCCAGAATGAAAAACCTGTCTGGGGCCCCCTCCCTGAGGCTACA
3401 3450

GTCGCCAAGGGGCAAGTTGGACTGGATTCCCAGCAGCCCCTCCCACTCCG
3451 3500

45 AGACAAAATCAGCTACCCTGGGGCAGGCCTCATTGGCCCCAGGAAACCCC
3501 3550

AGCCTGTCAGCACCTGTTCCAGGATCCAGTCCCAGCGCAGTA
3551

50 3592

AFP-TRE. SEQ ID NO:___.

1 GCATTGCTGTGAACTCTGTACTTAGGACTAACTTTGAGCAATAACACACATAGATTGAG
5 61 GATTGTTTGCTGTTAGCATACAACTCTGGTTCAAAGCTCCTCTTTATTGCTTGTCTTGG
121 AAAATTTGCTGTTCTTCATGGTTTCTCTTTTCACTGCTATCTATTTTCTCAACCACTCA
10 181 CATGGCTACAATAACTGTCTGCAAGCTTATGATTCCCAAATATCTATCTCTAGCCTCAAT
241 CTTGTTCCAGAAGATAAAAAGTAGTATTCAAATGCACATCAACGTCTCCACTTGGAGGGC
15 301 TTAAAGACGTTTCAACATACAAACCGGGGAGTTTTGCCTGGAATGTTTCCTAAAAATGTGT
361 CCTGTAGCACATAGGGTCCTCTTGTTCTTAAATCTAATTACTTTTAGCCCAGTGCTCA
421 TCCCACCTATGGGGAGATGAGAGTGAAAAGGGAGCCTGATTAATAATTACACTAAGTCAA
20 481 TAGGCATAGAGCCAGGACTGTTTGGGTAACTGGTCACTTTATCTTAACTAAATATATC
541 CAAAACCTGAACATGTACTTAGTTACTAAGTCTTTGACTTTATCTCATTCATACCACTCAG
25 601 CTTTATCCAGGCCACTTATGAGCTCTGTGTCCTTGAACATAAAATACAAATAACCGCTAT
661 GCTGTTAATTATTGGCAAATGTCCCATTTTCAACCTAAGGAAATACCATAAAGTAACAGA
721 TATACCAACAAAAGGTTACTAGTTAACAGGCATTGCCTGAAAAGAGTATAAAAGAATTTTC
30 781 AGCATGATTTTCCATATTGTGCTTCCACCACTGCCAATAACA (822)

Probasin -TRE SEQ ID NO:___

-426
5' -AAGCTTCCACAAGTGCATTTAGCCTCTCCAGTATTGCTGATGAATCCACAGT
40 TCAGGTTCAATGGCGTTCAAACCTTGATCAAAAATGACCAGACTTTATATTTA
CACCAACATCTATCTGATTGGAGGAATGGATAATAGTCATCATGTTTAAACAT
45 CTACCATTCCAGTTAAGAAAATATGATAGCATCTTGTTCTTAGTCTTTTCTTA
ARE-1
ATAGGGACATAAAGCCCACAAATAAAAATATGCCTGAAGAATGGGACAGGC
50 ATTGGGCATTGTCCATGCCTAGTAAAGTACTCCAAGAACCCTATTTGTATACTA
ARE-2
GATGACACAATGTCAATGTCTGTGTACAACGCCAACTGGGATGCAAGACAC
55 TGCCCATGCCAATCATCCTGAAAAGCAGCTATATAAAGCAGGAAGCTACTCT
CAAT box TATAA box

+1

+28

GCACCTTGTCTAGGTCCAGATACCTACAG-3'
Transcription site

5

Tyrosinase-TRE SEQ ID NO: __

PinAl end -1956

1 CCGGTGAAAATGATAAGTTGAATTCTGTCTTCGAGAACATAGAAAAGAA
51 TTATGAAATGCCAACATGTGGTTACAAGTAATGCAGACCCAAGGCTCCCC
10 101 AGGGACAAGAAGTCTTGTGTTAACTCTTGTGGCTCTGAAAGAAAGAGAG
151 AGAGAAAAGATTAAGCCTCCTTGTGGAGATCATGTGATGACTTCCTGATT
201 CCAGCCAGAGCGAGCATTTCATGGAACTTCTCTTCTCTTCCTCGAG -1716
-231 251 ATTACTAACCTTATTGTTAATATTCTAACCATAAGAATTAACCTATTAAT
301 GGTGAATAGAGTTTTTCACTTTAACATAGGCCTATCCCACTGGTGGGATA
15 351 CGAGCCAATTTCGAAAGAAAAGTCAGTCATGTGCTTTTCAGAGGATGAAA
401 GCTTAAGATAAAGACTAAAAGTGTGTTGATGCTGGAGGTGGGAGTGGTATT
451 ATATAGGTCTCAGCCAAGACATGTGATAATCACTGTAGTAGTAGCTGGAA
501 AGAGAAATCTGTGACTCCAATTAGCCAGTTCCTGCAGACCTTGTGA + 65

PinAl end

20

Human glandular kallikrein-TRE SEQ ID NO: __

gaattcagaa ataggggaag gttgaggaag gacactgaac tcaaagggga tacagtgatt 60
ggtttatttg tcttctcttc acaacattgg tgctggagga attcccaccc tgagggttatg
25 120
aagatgtctg aacacccaac acatagcact ggagatatga gctcgacaag agtttctcag
180
30 ccacagagat tcacagccta gggcaggagg aactgtacg ccaggcagaa tgacatggga
240
attgcgctca cgattggcctt gaagaagcaa ggactgtggg aggtgggctt tgtagtaaca
300
35 agagggcagg gtgaactctg attcccatgg gggaatgtga tggcctggtt acaaattttt
360
40 caagctggca gggaataaaa cccattacgg tgaggacctg tggagggcgg ctgccccaac
420
tgataaagga aatagccagg tgggggcctt tcccattgta ggggggacat atctggcaat
480
45 agaagccttt gagacccttt aggggtacaag tactgaggca gcaaataaaa tgaaatctta
540
tttttcaact ttatactgca tgggtgtgaa gatataattg tttctgtaca gggggtgagg
600

50

gaaaggaggg gaggaggaaa gttcctgcag gtctggtttg gtcttgat ccagggggtc
660

5 ttggaactat ttaaattaaa ttaaattaaa acaagcgact gttttaaatt aaattaaatt
720

aaattaaatt ttactttatt ttatcttaag ttctgggcta catgtgcagg acgtgcagct
780

10 ttgttacata ggtaaacgtg tgccatgggtg gtttgctgta cctatcaacc catcacctag
840

gtattaagcc cagcatgcat tagctgtttt tcttgacgct ctccctctcc ctgactccca
900

15 caacaggccc cagtgtgtgt tgttccccct cctgtgtcca tgtgttctca ttgttcagct
960

20 cccacttata agtgagaaca tgtgggtgtt ggttttctgt ttctgtgtta gtttgctgag
1020

gataatggct tccacctcca tccatgttcc tgcaaaggac gtgatcttat tcttttttat
1080

25 ggttgcatag aaattgtttt taaaaatcca attgatattg tatttaatta caagttaatc
1140

taattagcat actagaagag attacagaag atattaggta cattgaatga ggaaatatat
1200

30 aaaataggac gaaggtgaaa tattaggtag gaaaagtata atagttgaaa gaagtaaaaa
1260

35 aaaatatgca tgagtagcag aatgtaaaag aggtgaagaa cgtaatagtg actttttaga
1320

ccagattgaa ggacagagac agaaaaattt taaggaattg ctaaaccatg tgagtgttag
1380

40 aagtacagtc aataacatta aagcctcagg aggagaaaag aataggaaag gaggaatat
1440

gtgaataaat agtagagaca tgtttgatgg attttaaaat atttgaaaga cctcacatca
1500

45 aaggattcat accgtgccat tgaagaggaa gatggaaaag ccaagaagcc agatgaaagt
1560

50 tagaaatatt attggcaaag cttaaattgt aaaagtccta gagagaaagg atggcagaaa
1620

tattggcggg aaagaatgca gaacctagaa tataaattca tccaacagt ttggtagtgt
1680

55 gcagctgtag ccttttctag ataatacact attgtcatc atcgcttaag cgagtgtaaa
1740

atgggtctcct cactttatatt atttatatat ttatttagtt ttgagatgga gcctcgctct
1800

60

gtctcctagg ctggagtgca atagtgcgat accactcact gcaacctctg cctcctctgt
1860

5 tcaagtgatt ttcttacctc agcctcccgga gtagctggga ttacaggtgc gtgccaccac
1920

acccggctaa tttttgtatt tttttagag acggggtttt gccatgttgg ccaggtggt
1980

10 cttgaactcc tgacatcagg tgatccacct gccttggcct cctaaagtgc tgggattaca
2040

ggcatgagcc accgtgccc accactttat ttatttttta tttttatttt taaatttcag
2100

15 cttctatttg aaatacaggg ggcacatata taggattggt acatgggtat attgaactca
2160

20 ggtagtgatc atactaccca acaggtaggt tttcaaccca ctccccctct tttcctcccc
2220

attctagtag tgtgcagtgt ctattgttct catgtttatg tctatgtgtg ctccaggttt
2280

25 agtcccacc tgtaagtgag aacgtgtggt atttgatttt ctgtccctgt gtaattcac
2340

ttaggattat ggcttccagc tccattcata ttgctgtaaa ggatatgatt catttttcat
2400

30 ggccatgcag tattccatat tgcgtataga tcacattttc tttctttttt ttttttgaga
2460

35 cggagtcttg ctttgctgcc taggctggag tgcagtagca cgatctcggc tcaactgcaag
2520

cttcacctcc ggggttcacg tcattcttct gtctcagctt cccaagtagc tgggactaca
2580

40 ggcgcccgcc accacgtccg gctaattttt ttgtgtgttt ttagtagaga tgggggtttc
2640

actgtgttag ccaggatggt cttgatctcc tgaccttggt gtccacctgc ctcggtctcc
2700

45 caaagtgctg ggattacagg ggtgagccac tgcgcccggc ccatatatac cacattttct
2760

50 ttaaccaatc caccattgat gggcaactag gtagattcca tggattccac agttttgcta
2820

ttgtgtgcag tgtggcagta gacatatgaa tgaatgtgtc tttttggat aatgatttgc
2880

55 attccttttg gtatacagtc attaatagga gtgctgggtt gaacgggtggc tctgtttaaa
2940

attcctttgag aattttccaa actgtttgcc atagagagca aactaattta catttccacg
3000

60

aacagtatat aagcattccc ttttctccac agctttgtca tcatggtttt tttttttctt
3060

5

tatttttaaaa aagaatatgt tgttgttttc ccagggtaca tgtgcaggat gtgcaggttt
3120

gttacatagg tagtaaactg gagccatggg ggtttgctgc acctgtcaac ccattacctg
3180

10

ggtatgaagc cctgcctgca ttagctcttt tccctaatagc tctcactact gccccaccct
3240

caccctgaca gggcaaacag acaacctaca gaatgggagg aaatttttgc aatctattca
3300

15

tctgacaaaag gtcaagaata tccagaatct acaaggaact taagcaaatt tttacttttt
3360

20

aataatagcc actctgactg gcgtgaaatg gtatctcatt gtggttttca tttgaatttc
3420

tctgatgac agtgacgatg agcatttttt catatttggt ggctgcttgt acgtcttttg
3480

25

agaagtgtct cttcatgcct tttggccact ttaatgggat tattttttgc ttttttagttt
3540

aagttcctta tagattctgg atattagact tcttattgga tgcatagttt gtgaatactc
3600

30

tcttccattc tgtaggttgt ctgtttactc tattgatggc ttcttttgct gtgccgaagc
3660

35

atcttagttt aattagaaac cacctgccaa tttttgtttt tgttgcaatt gcttttgggg
3720

acttagtcat aaactctttg ccaaggctctg ggtcaagaag agtatttcct aggttttctt
3780

40

ctagaatttt gaaagtctga atgtaaacad ttgcattttt aatgcatctt gagttagttt
3840

ttgtatatgt gaaaggctca ctctcatttt ctttccctct ttctttcttt ctttcttttc
3900

45

tttctttctt tctttctttc tttctttctt tctttctttc tttctttttg tcttcttttc
3960

50

tttctttctt tcttctttct tctctctttc tttttttttt ttgatggagt attgctctgt
4020

tgcccaggct gcagtgcagc ggcacgatct cggctcactg caacctctgc ctctggggtt
4080

55

caactgattc tctgcatca gccttccaag tagctgggat tataggcgcc cgccaccacg
4140

cccgaactaat ttttgtattt ttagtagaga cgggggttgt ccatgttggc caggctgggt
4200

60

tgaaactcct gacctcaaac gatctgcctg ccttggcctc ccaaagtgcg gggattacag
 4260
 gtgtgagcca ctgtgccag ccaagaatgt cattttctaa gaggtccaag aacctcaaga
 4320
 tattttggga ccttgagaag agaggaattc atacagggtat tacaagcaca gcctaattggc
 4380
 aaatcttttg catggcttgg cttcaagact ttaggctctt aaaagtcgaa tccaaaaatt
 4440
 ttataaaaag ctccagctaa gctaccttaa aaggggcctg tatggctgat cactcttctt
 4500
 gctatacttt acacaaataa acaggccaaa tataatgagg ccaaaattta ttttgcaaat
 4560
 aaattggtcc tgctatgatt tactcttggg aagaacaggg aaaatagaga aaaatttaga
 4620
 ttgcatctga cctttttttc tgaattttta tatgtgccta caatttgagc taaatcctga
 4680
 attattttct gggtgcaaaa actctctaaa gaagaacttg gttttcattg tcttcgtgac
 4740
 acatttatct ggctctttac tagaacagct ttcttgtttt tgggtgttcta gcttggtgac
 4800
 cttacagttc tactcttcaa attattgta tgtgtatctc atagttttcc ttcttttgag
 4860
 aaaactgaag ccatgggtatt ctgaggacta gagatgactc aacagagctg gtgaatctcc
 4920
 tcatatgcaa tccactgggc tcgatctgct tcaaattgct gatgcactgc tgctaaagct
 4980
 atacatttaa aaccctcact aaaggatcag ggaccatcat ggaagaggag gaaacatgaa
 5040
 attgtaagag ccagattcgg ggggtagagt gtggagggtca gagcaactcc accttgaata
 5100
 agaaggtaaa gcaacctatc ctgaaagcta acctgccatg gtggcttctg attaacctct
 5160
 gttctaggaa gactgacagt ttgggtctgt gtcattgccc aaatctcatg ttaaattgta
 5220
 atccccagtg ttccggaggtg ggacttggtg gtaggtgatt cggatcatggg agtagatttt
 5280
 cttcttttggt gtgttacagt gatagtgagt gagttctcgt gagatctggt catttaaaag
 5340
 tgtgtggccc ctccccctccc tctcttggtc ctctactgc catgtaagat acctgtcct
 5400

gctttgcctt ctaccataag taaaagcccc ctgaggcctc cccagaagca gatgccacca
5460

5 tgccttctgt acagcctgca gaaccatcag ccaattaaac ctcttttctg tataaattac
5520

cagtcttgag tatctcttta cagcagtgtg agaacggact aatacaaggg tctccaaaat
5580

10 tccaagttta tgtattcttt cttgccaaat agcaggtatt taccataaat cctgtcctta
5640

ggatcaacaa ccttgatggc atcgacttc aattgtctta cacattcctt ctgaatgact
5700

15 cctccccat ggcataaag ccctgggtct tgggggataa tggcagaggg gtccaccatc
5760

20 ttgtctggct gccacctgag acacggacat ggcttctgtt ggtaagtctc tattaaatgt
5820

ttctttctaa gaaactggat ttgtcagctt gtttcttgg cctctcagct tcctcagact
5880

25 ttggggtagg ttgcacaacc ctgcccacca cgaacaaaat gtttaatatg ataaatatgg
5940

atagatataa tccacataaa taaaagctct tggagggccc tcaataattg ttaagagtgt
6000

30 aaatgtgtcc aaagatggaa aatgtttgag aactactgtc ccagagattt tcctgagttc
6060

35 tagagtgtgg gaatatagaa cctggagctt ggcttcttca gcctagaatc aggagtatgg
6120

ggctgaagtc tgaagcttgg cttcagcagt ttggggttgg cttccggagc acatatttga
6180

40 catgttgcca ctgtgatttg gggtttgga tttgctctga atcctaattg ctgtccttga
6240

ggcatctaga atctgaaatc tgtggtcaga attctattat cttgagtagg acatctccag
6300

45 tcctggttct gccttctagg gctggagtct gtagtcagtg acccgtctg gcatttcaac
6360

50 ttcatataca gtgggctatc ttttgggtcca tgtttcaacc aaacaaccga ataaaccatt
6420

agaacctttc cccacttccc tagctgcaat gttaaacctc ggatttctgt ttaatagggt
6480

55 catatgaata atttcagcct gatccaactt tacattcctt ctaccgttat tctacaccca
6540

ccttaaaaat gcattcccaa tatattccct ggattctacc tatatatggg aatcctgggt
6600

ttgccagttt ctagtgcatt aacatacctg atttacattc ttttacttta aagtggaaat
6660

5 aagagtcctt ctgcagagtt caggagttct caagatggcc cttacttctg acatcaattg
6720

agatttcaag ggagtcgcca agatcatcct caggttcagt gattgctggt agccctcata
6780

10 taactcaatg aaagctgtta tgctcatggc tatggtttat tacagcaaaa gaatagagat
6840

gaaaatctag caagggaaga gttgcatggg gcaaagacaa ggagagctcc aagtgcagag
6900

15 attcctgttg ttttctccca gtggtgtcat ggaaagcagt atcttctcca tacaatgatg
6960

20 tgtgataata ttcagtgtat tgccaatcag ggaactcaac tgagccttga ttatattgga
7020

gcttggttgc acagacatgt cgaccacctt catggctgaa ctttagtact tagccctcc
7080

25 agacgtctac agctgatagg ctgtaacca acattgtcac cataaatcac attgttagac
7140

tatccagtgt ggcccaagct cccgtgtaaa cacaggcact ctaaacaggc aggatatttc
7200

30 aaaagcttag agatgacctc ccaggagctg aatgcaaaga cctggcctct ttgggcaagg
7260

35 agaatccttt accgcacact ctcttcaca gggttattgt gaggatcaaa cgtggtcatg
7320

tgtgtgagac accagcacat gtctggctgt ggagagtac ttctatgtgt gctaacattg
7380

40 ctgagtgcta agaaagtatt aggcattggc ttcagcactc acagatgctc atctaactct
7440

cacaacatgg ctacaggggtg ggcactacta gcctcatttg acagaggaaa ggactgtgga
7500

45 taagaagggg gtgaccaata ggtcagagtc attctggatg caaggggctc cagaggacca
7560

50 tgattagaca ttgtctgcag agaaattatg gctggatgtc tctgccccgg aaagggggat
7620

gcactttcct tgaccccta tctcagatct tgactttgag gttatctcag acttcctcta
7680

55 tgataccagg agcccatcat aatctctctg tgcctctcc ccttcctcag tcttactgcc
7740

cactcttccc agctccatct ccagctggcc aggtgtagcc acagtaccta actctttgca
7800

60

gagaactata aatgtgtatc ctacagggga gaaaaaaaaa aagaactctg aaagagctga
7860

5 cattttaccg acttgcaaac acataagcta acctgccagt tttgtgctgg tagaactcat
7920

gagactcctg ggtcagaggc aaaagatttt attaccaca gctaaggagg cagcatgaac
7980

10 tttgtgttca catttgttca ctttgcccc caattcatat gggatgatca gagcagttca
8040

ggatgatgga cacaggggtt tgtggcaaag gtgagcaacc taggcttaga aatcctcaat
8100

15 cttataagaa ggtactagca aacttgtcca gtctttgtat ctgacggaga tattatcttt
8160

20 ataattgggt tgaaagcaga cctactctgg aggaacatat tgtatttatt gtctgaaca
8220

gtaaacaaat ctgctgtaaa atagacgtta actttattat ctaaggcagt aagcaaacct
8280

25 agatctgaag gcgataccat cttgcaaggc tatctgctgt acaaataatgc ttgaaaagat
8340

gggccagaaa agaaaacggt attattgcct ttgctcagaa gacacacaga aacataagag
8400

30 aaccatggaa aattgtctcc caacactgtt caccagagc cttccactct tgtctgcagg
8460

35 acagtcttaa catcccatca ttagtgtgtc taccacatct ggcttcaccg tgccctaacca
8520

agatttctag gtccagttcc ccaccatgtt tggcagtgcc cactgcca cccagaata
8580

40 agggagtgtc cagaattccg aggggacatg ggtggggatc agaacttctg ggcttgagtg
8640

cagagggggc ccatactcct tggttccgaa ggaggaagag gctggaggtg aatgtccttg
8700

45 gaggggagga atgtgggttc tgaactotta aatccccaag ggaggagact ggtaagggtc
8760

50 cagcttccga ggtactgacg tgggaatggc ctgagaggtc taagaatccc gtatcctcgg
8820

gaaggagggg ctgaaattgt gaggggttga gttgcagggg tttgttagct tgagactcct
8880

55 tgggtgggtc ctgggaagca aggactggaa ccattggctc cagggtttgg tgtgaaggta
8940

atgggatctc ctgattctca aagggtcaga ggactgagag ttgcccatgc tttgatcttt
9000

60

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

ccattctactc cttactccac ttgagggtaa tcacctactc ttctagttcc acaagagtgc
 9060
 gcctgcgcga gtataatctg cacatgtgcc atgtcccagag gcctggggca tcatccactc
 9120
 atcattcagc atctgcgcta tgcgggcgag gccggcgcca tgacgtcatg tagctgcgac
 9180
 tatccctgca gcgcgcctct cccgtcacgt cccaaccatg gagctgtgga cgtgcgctcc
 9240
 ctgggtggatg tggcctgcgt ggtgccagge cggggcctgg tgtccgataa agatcctaga
 9300
 accacaggaa accaggactg aaaggtgcta gagaatggcc atatgtcgct gtccatgaaa
 9360
 tctcaaggac ttctgggtgg agggcacagg agcctgaact tacgggtttg cccagttcca
 9420
 ctgtcctccc aagtgagtct cccagatacg aggcactgtg ccagcatcag cttcatctgt
 9480
 accacatctt gtaacagga ctaccagga ccctgatgaa caccatggtg tgtgcaggaa
 9540
 gagggggtga aggcattggac tcctgtgtgg tcagagccca gagggggcca tgacgggtgg
 9600
 ggaggaggct gtggactggc tcgagaagtg ggatgtggtt gtgtttgatt tcctttggcc
 9660
 agataaagtg ctggatatag cattgaaaac ggagtatgaa gaccagttag aatggagggt
 9720
 caggttggag ttgagttaca gatggggtaa aattctgctt cggatgagtt tggggattgg
 9780
 caatctaaag gtggtttggg atggcatggc tttgggatgg aaataggttt gtttttatgt
 9840
 tggctgggaa ggggtgtggg attgaattgg ggatgaagta ggtttagttt tggagataga
 9900
 atacatggag ctggctattg catgcgagga tgtgcattag tttggtttga tctttaaata
 9960
 aaggaggcta ttagggttgt cttgaattag attaagttgt gttgggttga tgggttgggc
 10020
 ttgtgggtga tgtgggttga ttgggctgtg ttaaattggt ttgggtcagg ttttgggtga
 10080
 ggttatcatg gggatgagga tatgcttggg acatggattc aggtggttct cattcaagct
 10140
 gaggcaaatt tcctttcaga cggtcattcc agggaaacgag tggttgtgtg ggggaaatca
 10200

ggccactggc tgtgaatata cctctatcct ggtcttgaat tgtgattata tatgtccatt
10260

5

ctgtctcctt cactgtactt ggaattgata tggtcattca gctggaaatg ggggaagatt
10320

ttgtcaaatt cttgagacac agctgggtct ggatcagcgt aagccttcct tctggtttta
10380

10

ttgaacagat gaaatcacat tttttttttc aaaatcacag aaatcttata gagttaacag
10440

tggactctta taataagagt taacaccagg actcttattc ttgattcttt tctgagacac
10500

15

caaaatgaga tttctcaatg ccaccctaata tctttttttt tttttttttt tttttgagac
10560

20

acagtctggg tcttttgctc tgtcactcag gctggagcgc agtgggtgtga tcatagctca
10620

ctgaaccctt gacctcctgg acttaaggga tctcctctgt tcagcctcct gagtagatgg
10680

25

ggctacaggt gcttgccacc acacctggct aattaaattt tttttttttt tttgtagaga
10740

aagggtctca ctttggtgcc ctggctgata ttgaacttct gacttcaagt gattcttcag
10800

30

ccttggtctc ccaaagcact gggattgctg gcatgagcca ctcaccgtgc ctggcttgca
10860

35

gcttaatctt ggagtgtata aacctggctc ctgatagcta gacatttcag tgagaaggag
10920

gcattggatt ttgcatgagg acaattctga cctaggaggg caggtcaaca ggaatccccg
10980

40

ctgtacctgt acgttggtaca ggcattggaga atgaggagtg aggaggccgt accggaaccc
11040

catattgttt agtggacatt ggattttgaa ataataggga acttgggtctg ggagagtcatt
11100

45

atttctggat tggacaatat gtggtatcac aagggtttat gatgaggagg aaatgtatgt
11160

50

ggggaaccat tttctgagtg tggaagtgc agaatacagag agtagctgaa tgccaacgct
11220

tctatttcag gaacatggta agttggagggt ccagctctcg ggctcagacg ggtataggga
11280

55

ccaggaagtc tcacaatccg atcattctga tatttcaggg catattagggt ttgggggtgca
11340

aaggaagtac ttgggactta ggcacatgag actttgtatt gaaaatcaat gattggggct
11400

60

ggccgtggtg ctcacgcctg taatctcatc actttgggag accgaagtgg gaggatggct
11460

5

tgatctcaag agttggacac cagcctaggg aacatggcca gacctctct ctacaaaaaa
11520

attaaaaatt agctggatgt ggtggtgcat gcttggtggtc tcagctatcc tggaggctga
11580

10

gacaggagaa tcggttgagt ctgggagttc aaggctacag ggagctgcga tcacgccgct
11640

gcactccagc ctgggaaaca gagtgcgact gtctcagaat ttttttaaaa aagaatcagt
11700

15

gatcatccca acccctgttg ctgttcatcc tgagcctgcc ttctctggct ttgttcccta
11760

20

gatcacatct ccatgatcca taggcctgc ccaatctgac ctacacccgt gggaaatgcct
11820

ccagactgat ctagtatgtg tggaacagca agtgctggct ctccctcccc ttccacagt
11880

25

ctgggtgtgg gagggggttg tccagcctcc agcagcatgg ggagggcctt ggtcagcatc
11940

taggtgccaa cagggcaagg gcggggctct ggagaatgaa ggctttatag ggctcctcag
12000

30

ggaggccccc cagccccaaa ctgcaccacc tggccgtgga caccggt
12047

35

HRE-TRE SEQ ID NO:___

ccccgagg cagtgcag gaggtcagg gcgtgcgt ggtcgcagcgagaccccg gggtgcag

gccgga

40

PSA-TRE SEQ ID NO:___

aagcttctag ttttcttttc ccggtgacat cgtggaaagc actagcatct ctaagcaatg 60

45

atctgtgaca atattcacag tgtaatgcc tccagggaac tcaactgagc cttgatgtcc
120

agagattttt gtgttttttt ctgagactga gtctcgctct gtgccaggct ggagtgcagt
180

50

ggtgcaacct tggctcactg caagctccgc ctctgggtt caccgccattc tctgcctca
240

55

gcctcctgag tagctgggac tacaggcacc cgccaccacg cctggctaata ttttttgtat
300

ttttagtaga gatgggggttt cactgtgtta gccaggatgg tctcagtctc ctgacctcgt
360

5

gatctgcccc ccttggcctc ccaaagtgtt gggatgacag gcgtgagcca ccgcgcctgg
420

ccgatatcca gagatttttt ggggggctcc atcacacaga catgttgact gtcttcatgg
480

10

ttgactttta gtatccagcc cctctagaaa tctagctgat atagtgtggc tcaaaacctt
540

cagcacaaat cacaccgtta gactatctgg tgtggcccaa accttcaggt gaacaaaggg
600

15

actctaattt ggcaggatac tccaaagcat tagagatgac ctcttgcaaa gaaaaagaaa
660

20

tggaaaagaa aaagaaagaa aggaaaaaaa aaaaaaaaaa gagatgacct ctcaggctct
720

gaggggaaac gcctgaggtc tttgagcaag gtcagtcctc tgttgcacag tctccctcac
780

25

agggtcattg tgacgatcaa atgtgggtcac gtgtatgagg caccagcaca tgccctggctc
840

tggggagtgc cgtgtaagtg tatgcttgca ctgctgaatg gctgggatgt gtcagggatt
900

30

atcttcagca cttacagatg ctcatctcat cctcacagca tcactatggg atgggtatta
960

ctggcctcat ttgatggaga aagtggctgt ggctcagaaa ggggggacca ctagaccagg
1020

35

gacactctgg atgctgggga ctccagagac catgaccact caccaactgc agagaaatta
1080

40

attgtggcct gatgtccctg tcttgagagag ggtggaggtg gaccttcact aacctoctac
1140

cttgacctc tcttttaggg ctctttctga cctccaccat ggtactagga cccattgta
1200

45

ttctgtacc ccttgactct atgaccccca ccgccactg catccagctg ggtccctcc
1260

50

tatctctatt cccagctggc cagtgcagtc tcagtgccca cctgtttgtc agtaactctg
1320

aaggggctga cattttactg acttgcaaac aaataagcta actttccaga gttttgtgaa
1380

55

tgctggcaga gtccatgaga ctctgagtc agaggcaaag gcttttactg ctcacagctt
1440

agcagacagc atgaggttca tgttcacatt agtacacctt gccccccca aatctttag
1500

60

ggtgaccaga gcagtctagg tggatgctgt gcagaagggg tttgtgccac tggtgagaaa
1560

5

cctgagatta ggaatcctca atcttatact gggacaactt gcaaactgc tcagcctttg
1620

tctctgatga agatattatc ttcattgatct tggattgaaa acagacctac tctggaggaa
1680

10

catattgtat cgattgtcct tgacagtaaa caaatctgtt gtaagagaca ttatctttat
1740

tatctaggac agtaagcaag cctggatctg agagagatat catcttgcaa ggatgcctgc
1800

15

tttacaacaa tccttgaaac aacaatccag aaaaaaaaaag gtgttactgt ctttgctcag
1860

20

aagacacaca gatactgac agaaccatgg agaattgcct cccaacgctg ttcagccaga
1920

gccttccacc ctttctgcag gacagtctca acgttccacc attaaatact tcttctatca
1980

25

catcccgttt ctttatgcct aaccaagggt ctaggtcccg atcgactgtg tctggcagca
2040

ctccactgcc aaaccagaa taaggcagcg ctcaggatcc cgaaggggca tggctgggga
2100

30

tcagaacttc tgggtttgag tgaggagtgg gtccaccctc ttgaatttca aaggaggaag
2160

35

aggctggatg tgaagggtact gggggaggga aagtgtcagt tccgaactct taggtcaatg
2220

agggaggaga ctggtaagggt ccagctccc gaggtactga tgtgggaatg gcctaagaat
2280

40

ctcatatcct caggaagaag gtgctggaat cctgaggggt agagttctgg gtatatattgt
2340

ggcttaaggc tctttggccc ctgaaggcag aggctggaac cattaggtcc agggtttggg
2400

45

gtgatagtaa tgggatctct tgattcctca agagtctgag gatcgagggt tgcccattct
2460

50

tccatcttgc cacctaattc ttactccact tgagggtatc accagccctt ctagtccat
2520

gaagggtccc tgggcaagca caatctgagc atgaaagatg cccagaggc cttgggtgtc
2580

55

atccactcat catccagcat cacactctga ggggtgtggc agcaccatga cgtcatgttg
2640

ctgtgactat ccctgcagcg tgcctctcca gccacctgcc aaccgtagag ctgcccatcc
2700

60

tcctctggtg ggagtggcct gcatggtgcc aggctgaggc ctagtgtcag acaggagacc
2760

5

tggaatcata gggatccagg actcaaaagt gctagagaat ggccatatgt caccatccat
2820

gaaatctcaa gggcttcttg gtggagggca caggacactg aacttatggt ttcccaagtc
2880

10

tattgctctc ccaagtgagt ctcccagata cgaggcactg tgccagcatc agccttatct
2940

ccaccacatc ttgtaaaagg actaccagg gccctgatga acaccatggt gtgtacagga
3000

15

gtagggggtg gaggcacgga ctctgtgag gtcacagcca agggagcatc atcatgggtg
3060

20

gggaggaggc aatggacagg cttgagaacg gggatgtggt tgtatttggt tttctttggt
3120

tagataaagt gctgggtata ggattgagag tggagtatga agaccagtta ggatggagga
3180

25

tcagattgga gttgggttag ataaagtgct gggatatagga ttgagagtgg agtatgaaga
3240

ccagttagga tggaggatca gattggagtt gggtagaga tggggtaaaa ttgtgctccg
3300

30

gatgagtttg ggattgacac tgtggaggtg gtttgggatg gcatggcttt gggatggaaa
3360

35

tagatttggt ttgatgttgg ctacagacatc cttggggatt gaactgggga tgaagctggg
3420

tttgattttg gaggtagaag acgtggaagt agctgtcaga ttgacagtg gccatgagtt
3480

40

ttgtttgatg gggaatcaaa caatggggga agacataagg gttggcttgt taggttaagt
3540

tgcgttgggt tgatggggtc ggggctgtgt ataatgcagt tggattggtt tgtattaaat
3600

45

tgggttgggt cagggttttg ttgaggatga gttgaggata tgcttgggga caccggatcc
3660

50

atgaggttct cactggagtg gagacaaact tcctttccag gatgaatcca gggaagcctt
3720

aattcacgtg taggggaggt caggccactg gctaagtata tccttccact ccagctctaa
3780

55

gatggtctta aattgtgatt atctatatcc acttctgtct cctcactgt gcttggagtt
3840

tacctgatca ctcaactaga aacaggggaa gattttatca aattcttttt tttttttttt
3900

60

tttttttgag acagagtctc actctgttgc ccaggctgga gtgcagtggc gcagtctcgg
3960

5

ctcactgcaa cctctgcctc ccagggtcaa gtgattctcc tgccctcagcc tccctgagttg
4020

ctgggattac aggcattgcag caccatgccc agctaatttt tgtattttta gtagagatgg
4080

10

ggtttcacca atgtttgccca ggctggcctc gaactcctga cctgggtgatc cacctgcctc
4140

agcctcccaa agtgcctggga ttacaggcgt cagccaccgc gccagccac ttttgtcaaa
4200

15

ttcttgagac acagctcggg ctggatcaag tgagctactc tggttttatt gaacagctga
4260

20

aataaccaac tttttggaaa ttgatgaaat cttacggagt taacagtggga ggtaccaggg
4320

ctcttaagag ttcccgatcc tcttctgaga ctacaaattg tgattttgca tgccacctta
4380

25

atcttttttt tttttttttt aaatcgaggt ttcagtctca ttctatttcc caggctggag
4440

ttcaatagcg tgatcacagc tcaactgtagc cttgaactcc tggccttaag agattctcct
4500

30

gcttcgggtct cccaatagct aagactacag tagtccacca ccatatccag ataattttta
4560

35

aatttttttg ggggccgggc acagtggctc acgcctgtaa tcccaacacc atgggaggct
4620

gagatgggtg gatcacgagg tcaggagttt gagaccagcc tgaccaacat ggtgaaactc
4680

40

tgtctctact aaaaaaaaaa aaaatagaaa aattagccgg gcgtgggtggc acacggcacc
4740

tgtaatccca gctactgagg aggctgaggc aggagaatca cttgaaccca gaaggcagag
4800

45

gttgcaatga gccgagattg cgccactgca ctccagcctg ggtgacagag tgagactctg
4860

50

tctcaaaaaa aaaaaatttt tttttttttt ttgtagagat ggatcttget ttgtttctct
4920

ggttggcctt gaactcctgg cttcaagtga tcctcctacc ttggcctcgg aaagtgttgg
4980

55

gattacaggc gtgagccacc atgactgacc tgtcgttaat cttgaggtac ataaacctgg
5040

ctcctaaagg ctaaaggcta aatatttggt ggagaagggg cattggattt tgcattgagga
5100

60

tgattctgac ctgggagggc aggtcagcag gcattctctgt tgcacagata gagtgtacag
5160

5

gtctggagaa caaggagtgg ggggttattg gaattccaca ttgtttgctg cacgttggat
5220

tttgaaatgc tagggaactt tgggagactc atatttctgg gctagaggat ctgtggacca
5280

10

caagatcttt ttatgatgac agtagcaatg tatctgtgga gctggattct gggttgggag
5340

15

tgcaaggaaa agaattgtact aaatgccaaag acatctattt caggagcatg aggaataaaa
5400

gttctagttt ctggtctcag agtgggtgcat ggatcagggg gtctcacaat ctctgagtg
5460

20

ctggtgtctt agggcacact ggggtcttga gtgcaaagga tctaggcacg tgaggctttg
5520

tatgaagaat cggggatcgt acccaccccc tgtttctgtt tcatcctggg catgtctcct
5580

25

ctgcctttgt ccctagatg aagtctccat gagctacaag gccctggtgc atccagggtg
5640

atctagtaat tgcagaacag caagtgctag ctctccctcc cctccacag ctctgggtgt
5700

30

gggagggggg tgtccagcct ccagcagcat ggggagggcc ttggtcagcc tctgggtgcc
5760

35

agcagggcag gggcggagtc ctggggaatg aaggttttat agggctcctg ggggaggctc
5820

cccagcccca agctt
5835

40

CEA TRE SEQ ID NO:

aagcttttta gtgctttaga cagtgaagctg gtctgtctaa cccaagtga cttgggtcca	60
tactcagccc cagaagtga ggggtgaagct ggggtggagcc aaaccaggca agcctaccct	120
cagggctccc agtggcctga gaaccattgg acccaggacc cattacttct agggtaagga	180
aggtacaaac accagatcca accatgggtct ggggggacag ctgtcaaatz cctaaaaata	240
tacctgggag aggagcaggc aaactatcac tgccccaggt tctctgaaca gaacagagg	300
ggcaacccaa agtccaaatc caggtgagca ggtgcaccaa atgcccagag atatgacgag	360
gcaagaagtg aaggaaccac ccctgcatca aatgttttgc atgggaagga gaaggggggt	420
gctcatgttc ccaatccagg agaatgcatt tgggatctgc cttcttctca ctcttggtt	480
agcaagacta agcaaccagg actctggatt tggggaaaga cgtttatttg tggaggccag	540
tgatgacaat cccacgaggg cctaggtgaa gagggcagga aggctcgaga cactggggac	600
tgagtgaaaa ccacacccat gatctgcacc acccatggat gctccttcat tgctcacctt	660
tctgttgata tcagatggcc ccattttctg taccttcaca gaaggacaca ggctagggtc	720
tgtgcatggc ctcatcccc ggggccatgt gaggacagca ggtgggaaag atcatgggtc	780
ctctgggtc ctgcagggcc agaacattca tcaccatac tgacctcta gatgggaatg	840
gcttccttg ggtgggcca acggggcctg ggcaggggag aaaggacgtc aggggacagg	900
gaggaagggt catcgagacc cagcctggaa ggttctgtc tctgaccatc caggatttac	960
ttcctgcat ctaccttggt tcattttccc tcagcaatga ccagctctgc ttctgatct	1020
cagcctccca ccctggacac agcaccocag tccctggccc ggctgcatcc acccaatacc	1080
ctgataaacc aggaccatt acttctaggg taaggagggt ccaggagaca gaagctgagg	1140
aaaggctctga agaagtcaca tctgtcctgg ccagagggga aaaaccatca gatgctgaac	1200
caggagaatg ttgaccagg aaagggaccg aggaccaag aaaggagtca gaccaccagg	1260
gtttgcctga gaggaaggat caaggccccg agggaaagca gggctggctg catgtgcagg	1320
acactgggtg ggcataatgt tcttagattc tccctgaatt cagtgtccct gccatggcca	1380
gactctctac tcaggcctgg acatgctgaa ataggacaat ggccttgtcc tctctccca	1440
ccatttgga agagacataa aggacattcc aggacatgcc ttctgggag gtccaggttc	1500
tctgtctcac acctcagga ctgtagttac tgcatcagcc atggtaggtg ctgatctcac	1560
ccagcctgtc caggcccttc cactctccac tttgtgacca tgtccaggac caccctcag	1620
atcctgagcc tgcaaatacc cccttgcctg gtgggtggat tcagtaaaca gtgagctcct	1680

atccagcccc-gagagccacc tctgtcacct tctgtctggg catcatccca ccttcacaag	1740
cactaaagag catggggaga cctggctagc tgggtttctg catcacaag aaaataatcc	1800
cccaggttcg gattcccagg gctctgtatg tggagctgac agacctgagg ccaggagata	1860
gcagaggtca gccctaggga ggggtgggtca tccaccagc ggacaggggt gcaccagcct	1920
tgctactgaa agggcctccc caggacagcg ccatcagccc tgcctgagag ctttgctaaa	1980
cagcagtcag aggaggccat ggcagtggct gagctcctgc tccaggcccc aacagaccag	2040
acceacagca caatgcagtc ctccccaac gtcacaggtc accaaaggga aactgaggtg	2100
ctacctaac ttagagccat caggggagat aacagccaa tttcccaaac aggccagttt	2160
caatcccatg acaatgacct ctctgtctc attcttccca aataggacg ctgattctcc	2220
cccaccatgg atttctcct tgtcccggga gccttttctg cccctatga tctgggcact	2280
cctgacacac acctcctctc tggtgacata tcagggtccc tactgtcaa gcagtcaga	2340
aaggacagaa ccttggacag cgccatctc agcttcaccc ttcctcctc acagggttca	2400
gggcaaagaa taaatggcag aggccagtga gccagagat ggtgacaggc agtgaccag	2460
gggcagatgc ctggagcagg agctggcggg gccacaggga gaagggtatg caggaaggga	2520
aaccagaaa tgggcaggaa aggaggacac aggtctgtg gggctgcagc ccagggttgg	2580
actatgagtg tgaagccatc tcagcaagta aggccaggtc ccatgaacaa gagtgggagc	2640
acgtggcttc ctgtctgtg tatgggggtg gggattccat gccccataga accagatggc	2700
cggggttcag atggagaagg agcaggacag gggatcccca ggataggagg acccagtg	2760
cccccccag gcaggtgact gatgaatgg catgcagggt cctcctgggc tgggctctcc	2820
ctttgtcct caggattcct tgaaggaaca tccggaagcc gaccacatct acctggtggg	2880
ttctggggag tccatgtaaa gccaggagct tgtgttgcta ggaggggtca tggcatgtgc	2940
tgggggcacc aaagagagaa acctgagggc aggcaggacc tggctctgagg aggcattgga	3000
gccagatgg ggagatggat gtcaggaaa gctgccccat caggagggt gatagcaatg	3060
gggggtctgt gggagtgggc acgtgggatt cctgggctc tgccaagttc cctcccatag	3120
tcacaacctg gggacactgc ccatgaagg gcgcctttgc ccagccagat gctgctggtt	3180
ctgcccattc actacctct ctgtccagc cactctgggt cttctccag atgccctgga	3240
cagcctggc ctgggcctgt cccctgagag gtgttgggag aagctgagtc tctggggaca	3300
ctctcatcag agtctgaaag gcacatcagg aaacatccct ggtctccagg actaggcaat	3360

gaggaaaggg ccccagctcc tccctttgcc actgagaggg togaccctgg gtggccacag	3420
tgactttctgc gtctgtccca gtcaccctga aaccacaaca aaaccccgag cccagaccct	3480
gcaggtacaa tacatgtggg gacagtctgt acccagggga agccagttct ctcttcctag	3540
gagaccgggc ctcaagggtg tgcccggggc aggcgggggc agcacgtgcc tgtccttgag	3600
aactcgggac cttaagggtc tctgctctgt gaggcacagc aaggatcctt ctgtccagag	3660
atgaaagcag ctctgcccc tctctgacc tcttctctct tcccaaactt caaccaacaa	3720
ataggtgttt caaatctcat catcaaactt tcatccatcc acatgagaaa gcttaaaacc	3780
caatggattg acaacatcaa gagttggaac aagtggacat ggagatgtta cttgtggaaa	3840
tttagatgtg ttcagctatc gggcaggaga atctgtgtca aattccagca tggttcagaa	3900
gaatcaaaaa gtgtcacagt ccaaagtgc aacagtgcag gggataaaac tgtggtgcat	3960
tcaaactgag ggatattttg gaacatgaga aaggaaggga ttgctgctgc acagaacatg	4020
gatgatctca cacatagagt tgaaagaaag gagtcaatcg cagaatagaa aatgatcact	4080
aattccacct ctataaagt tccaagagga aaaccaatt ctgctgctag agatcagaat	4140
ggaggtgacc tgtgccttgc aatggctgtg agggtcacgg gagtgtcact tagtgcaggc	4200
aatgtgccgt atcttaactt gggcagggtt ttcattgagca cataggaatg cagacattac	4260
tgtgtgttct attttacttc accggaaaag aagaataaaa tcagccgggc gcgggtggctc	4320
acgcctgtaa tcccagcact ttagaaggct gaggtgggca gattacttga ggtcaggagt	4380
tcaagaccac cctggccaat atggtgaaac cccggctcta ctaaaaatac aaaaattagc	4440
tgggcatggt ggtgcgcgcc tgtaatccca gctactcggg aggctgaggc tggacaattg	4500
cttggaacca ggaagcagag gttgcagtga gccaaagtgt tgccactgca ctccagcttg	4560
ggcaacagag ccagactctg taaaaaaaaa aaaaaaaaaa aaaaaaagaa agaaagaaaa	4620
agaaagaaaa gtataaaatc tctttgggtt aacaaaaaaaa gatccacaaa acaaacacca	4680
gctcttatca aacttacaca actctgccag agaacaggaa acacaaatac tcattaactc	4740
acttttgttg caataaaacc ttcattgtca aaggagacca ggacacaatg aggaagttaa	4800
actgcaggcc ctacttgggt gcagagaggg aaaatccaca aataaaacat taccagaagg	4860
agctaagatt tactgcattg agttcattcc ccaggtatgc aaggtgattt taacacctga	4920
aatcaatca ttgcctttac tacatagaca gattagctag aaaaaatta caactagcag	4980
aacagaagca atttggcctt cctaaaattc cacatcatat catcatgatg gagacagtgc	5040
agacgccaat gacaataaaa agagggacct ccgtcaccgg gtaaacatgt ccacacagct	5100

ccagcaagca-~~eccgtcttcc~~ cagtgaatca ctgtaacctc ccctttaatc agccccaggg 5160
aaggctgcct gcgatggcca cacaggctcc aaccogtggg cctcaacctc ccgcagaggg 5220
tctcctttgg ccaccccatg gggagagcat gaggacaggg cagagccctc tgatgcccac 5280
acatggcagg agctgacgcc agagccatgg gggctggaga gcagagctgc tggggtcaga 5340
gcttcctgag gacaccaggg cctaagggaa ggcagctccc tggatggggg caaccaggct 5400
ccgggctcca acctcagagc ccgcatggga ggagccagca ctctaggcct ttcctagggt 5460
gactctgagg ggaccctgac acgacaggat cgctgaatgc acccgagatg aagggggccac 5520
cacgggaccc tgctctcgtg gcagatcagg agagagtggg acaccatgcc agggcccccac 5580
ggcatggctg cgactgaccc agggcactcc cctgcatgca tcagcctcgg taagtcacat 5640
gaaccaagccc aggaccaatg tggaaggaag gaaacagcat cccctttagt gatggaaccc 5700
aaggctcagtg caaagagagg ccatgagcag ttaggaaggg tggccaacc tacagcacia 5760
accatcatct atcataagta gaagccctgc tccatgaccc ctgcatttaa ataaacgttt 5820
gttaaagtag tcaaattccc tcaccatgag agctcacctg tgtgtaggcc catcacacac 5880
acaaacacac acacacacac acacacacac acacacacac acagggaaag tgcaggatcc 5940
tgacagcac caggcaggct tcacaggcag agcaaacagc gtgaatgacc catgcagtgc 6000
cctgggcccc atcagctcag agaccctgtg agggctgaga tggggctagg caggggagag 6060
acttagagag ggtggggcct ccagggaggg ggctgcaggg agctgggtac tgcctccag 6120
ggagggggct gcaggagct ggggtactgcc ctccaggag ggggctgcag ggagctgggt 6180
actgccctcc agggaggggg ctgcaggag ctgggtactg cctccaggg agggggctgc 6240
aggagctgg gtactgcct ccagggaggc aggagcactg ttccaacag agagcacatc 6300
ttctgcagc agctgcacag acacaggagc cccatgact gccctgggcc aggggtgtga 6360
ttccaaattt cgtgccccat tgggtgggac ggaggtgac cgtgacatcc aaggggcac 6420
tgtgattoca aacttaaact actgtgcta caaatagga aataacctc cttttctac 6480
tatctcaaat tccctaagca caagctagca cctttaaat caggaagttc agtcactcct 6540
gggtcctcc catgccccca gtctgacttg cagggtgcaca gggtggtga catctgtcct 6600
tgctctcct cttggctcaa ctgccgcccc tctgggggt gactgatggt caggacaagg 6660
gatcctagag ctggcccat gattgacagg aaggcaggac ttggcctcca ttctgaagac 6720
taggggtgtc aagagagctg ggcattccac agagctgcac aagatgacgc ggacagaggg 6780

tgacacaggg ctcagggcct cagacgggtc gggagggctca gctgagagtt cagggacaga	6840
cctgaggagc ctcagtggga aaagaagcac tgaagtggga agttctggaa tgttctggac	6900
aagcctgagt gctctaagga aatgctccca cccgatgta gcctgcagca ctggacggtc	6960
tgtgtacctc cccgctgccc atcctctcac agcccccgcc tctagggaca caactcctgc	7020
cctaacatgc atctttcctg tctcattcca cacaaaaggg cctctggggg cctgtttctg	7080
cattgcaagg agtggaggtc acgttcccac agaccaccca gcaacagggg cctatggagg	7140
tgcggtcagg aggatcacac gtcccccat gccagggga ctgactctgg gggatgatga	7200
ttggcctgga ggccactggg cccctctgtc cctgagggga atctgcaccc tggaggctgc	7260
cacatccctc ctgattcttt cagctgaggg ccttcttga aatcccaggg aggactcaac	7320
cccactggg aaaggcccag tgtggacggg tccacagcag ccagctaa ggccttggac	7380
acagatcctg agtgagagaa cctttagga cacaggtgca cggccatgtc ccagtgccc	7440
acacagagca ggggcatctg gaccctgagt gtgtagctcc cgcgactgaa ccagccctt	7500
cccactgac gtgaccctg ggggtggctc aggtctccag tccatgccac caaatctcc	7560
agattgaggg tcctccctg agtccctgat gcctgtccag gagctgcccc ctgagcaaat	7620
ctagagtgca gagggctggg attgtggcag taaaagcagc cacatttgc tcaggaagga	7680
aagggaggac atgagctcca ggaagggcga tggcgtctc tagtgggcgc ctctgttaa	7740
tgagcaaaaa ggggccagga gagttgagag atcagggctg gccttggact aaggctcaga	7800
tggagaggac tgaggtgcaa agagggggct gaagtagggg agtggtcggg agagatggga	7860
ggagcaggta aggggaagcc ccagggaggc cgggggaggg tacagcagag ctctccactc	7920
ctcagcattg acatttgggg tggctgtgct agtgggggtc tgtaagttgt aggggttca	7980
gcaccatctg gggactctac ccactaatg ccagcaggac tcctcccca agctctaaca	8040
accaacaatg tctccagact ttccaatgt ccctggaga gcaaatgtc ttctggcaga	8100
atcactgatc tacgtcagtc tctaaaagt actcatcagc gaaatcctc acctcttggg	8160
agaagaatca caagtgtgag aggggtagaa actgcagact tcaaatctt tccaaaagag	8220
ttttacttaa tcagcagttt gatgtcccag gagaagatac atttagagtg ttagagtgtg	8280
atgccacatg gctgcctgta cctcacagca ggagcagagt gggttttcca agggcctgta	8340
accacaactg gaatgacact cactgggtta cattacaaag tggaatgtgg ggaattctgt	8400
agactttggg aagggaaatg tatgacgta gccacagcc taaggcagtg gacagtccac	8460
tttgaggctc tcaccatcta ggagacatct cagccatgaa catagccaca tctgtcatta	8520

gaaaacatgt-tttattaaga ggaaaaatct aggetagaag tgctttatgc tcttttttct	8580
ctttatgttc aaattcatat acttttagat cattcotta agaagaatct atccccctaa	8640
gtaaatgtta tcaactgactg gatagtgttg gtgtctcact cccaaccct gtgtggtgac	8700
agtgccctgc tccccagcc ctgggccctc tctgattcct gagagctttg ggtgctcctt	8760
cattaggagg aagagaggaa ggggtgtttt aatattctca ccattcacc atccacctct	8820
tagaactgga gaagaatcag ttgccactc ttggatttga tctcgaatt aatgacctct	8880
atttctgtcc cttgtccatt tcaacaatgt gacaggccta agaggctcct tctccatgtg	8940
atttttgagg agaaggttct caagataagt tttctcacac ctctttgaat tacctccacc	9000
tggtgcccc taccattac cagcagcatt tggacccttt ttctgttagt cagatgcttt	9060
ccacctcttg aggggtgtata ctgtatgctc tctacacagg aatatgcaga ggaaatagaa	9120
aaagggaaat cgcattacta ttcagagaga agaagacctt tatgtgaatg aatgagagtc	9180
caaaatccta agagagccca tataaaatta ttaccagtgc taaaactaca aaagttacac	9240
taacagtaaa ctagaataat aaaacatgca tcacagttgc tggtaagct aaatcagata	9300
cttttttctt agaaaaagca ttccatgtgt gttgcagtga tgacaggagt gcccttcagt	9360
caatatgctg cctgtaattt ttgttcctg gcagaatgta ttgtcttttc tccctttaa	9420
ctttaaatgc aaaactaaag gcagctcctg ggccccctcc ccaaagtcag ctgcctgcaa	9480
ccagccccac gaagagcaga ggctgagct tccctggtca aaataggggg ctaggagct	9540
taaccttgct cgataaagct gtgttccag aatgtcgtc ctgttccag gggcaccagc	9600
ctggaggggtg gtgagcctca ctggtggcct gatgcttacc ttgtgccctc acaccagtgg	9660
tcaactggaac cttgaacact tggctgtgc ccgcatctgc agatgtcaag aacttctgga	9720
agtcaaatta ctgccactt ctccaggga gatactgtg aacatccaaa accatgccac	9780
agaaccctgc ctgggtgcta caacacatat ggactgtgag caccaagtcc agccctgaat	9840
ctgtgaccac ctgccaagat gccctaact gggatccacc aatcactgca catggcaggc	9900
agcgaggctt ggaggtgctt cgccacaagg cagccccaat ttgtgggag tttcttgga	9960
cctggtagtg gtgaggagcc ttgggacct caggattact ccccttaagc atagtggga	10020
cccttctgca tccccagcag gtgccccgt cttcagagcc tctctctctg aggtttacc	10080
agaccctgc accaatgaga ccatgctgaa gcctcagaga gagagatgga gctttgacca	10140
ggagccgctc ttccttgagg gccagggcag ggaagcagg aggcagcacc aggagtggga	10200

acaccagtgt ctaagcccct gatgagaaca gpgtggtctc tcccatatgc ccataaccagg	10260
cctgtgaaca gaatcctcct tctgcagtga caatgtctga gaggacgaca tgtttcccag	10320
cctaacgtgc agccatgccc atctaccac tgcctactgc aggacagcac caaccaggga	10380
gctgggaagc tgggagaaga catggaatac ccatggcttc tcaccttcct ccagtccagt	10440
gggcaccatt tatgcctagg acaccacct gccggcccca ggctcttaag agttaggtca	10500
cctaggtgcc tctgggaggc cgaggcagga gaattgcttg aaccgggag gcagagggtg	10560
cagtgagccg agatcacacc actgcactcc agcctgggtg acagaatgag actctgtctc	10620
aaaaaaaaag agaaagatag catcagtggc taccaagggc taggggcagg ggaagggtga	10680
gagttaatga ttaatagtat gaagtttcta tgtgagatga tgaaaatgtt ctggaaaaaa	10740
aaatatagtg gtgaggatgt agaataattt gaataaatt aacggcattt aattgtacac	10800
ttaacatgat taatgtggca tattttatct tatgtatttg actacatcca agaaacactg	10860
ggagagggaa agcccaccat gtaaaataca cccaccctaa tcagatagtc ctcattgtac	10920
ccaggtacag gccctcatg acctgcacag gaataactaa ggatttaagg acatgaggct	10980
tcccagccaa ctgcagggtgc acaacataaa tgtatctgca aacagactga gagtaaagct	11040
gggggcacaa acctcagcac tgccaggaca cacacccttc tcgtggattc tgactttatc	11100
tgaccgggcc cactgtccag atcttgttgt gggattggga caaggagggt cataaagcct	11160
gtccccaggg cactctgtgt gagcacacga gacctccca cccccacc gttaggtctc	11220
cacacataga tctgaccatt aggcattgtg aggaggactc tagcgcgggc tcagggatca	11280
caccagagaa tcaggtagag agaggaagac ggggctcgag gagctgatgg atgacacaga	11340
gcagggttcc tgcagtcac aggtccagct caccctgggtg taggtgcccc atccccctga	11400
tccaggcatc cctgacacag ctccctcccg gagcctctc ccaggtgaca catcagggtc	11460
cctcactcaa gctgtccaga gagggcagca ccttgagacag cgcccacccc acttcactct	11520
tctccctca cagggtcag ggctcagggc tcaagtctca gaacaaatgg cagaggccag	11580
tgagcccaga gatggtgaca gggcaatgat ccaggggcag ctgcctgaaa cgggagcagg	11640
tgaagccaca gatgggagaa gatggttcag gaagaaaaat ccaggaatgg gcaggagagg	11700
agaggaggac acaggctctg tggggctgca gccaggatg ggactaagtg tgaagacatc	11760
tcagcaggtg aggccaggtc ccatgaacag agaagcagct cccacctccc ctgatgcacg	11820
gacacacaga gtgtgtggtg ctgtgcccc agagtcgggc tctcctgttc tgggtccccag	11880
ggagtgagaa gtgagggtga cttgtccctg ctctctctg ctaccccaac attcaccttc	11940

tectcatgcc-cctctctctc aaatatgatt tggatctatg tccccgccca aatctcatgt	12000
caaattgtaa accccaatgt tggaggtggg gccttgtgag aagtgattgg ataatgcggg	12060
tggattttct gctttgatgc tgtttctgtg atagagatct cacatgatct ggttgtttaa	12120
aagtgtgtag cacctctccc ctctctctct ctctctctta ctcatgctct gccatgtaag	12180
acgttctctgt tcccccttca ccgtccagaa tgattgtaag ttttctgagg cctccccagg	12240
agcagaagcc actatgcttc ctgtacaact gcagaatgat gagcgaatta aacctctttt	12300
ctttataaat taccagctct caggtatttc tttatagcaa tgcgaggaca gactaataca	12360
atcttctact ccagatccc cgcacacgt tagccccaga catcactgcc cctgggagca	12420
tgcacagcgc agcctctctgc cgacaaaagc aaagtcacaa aaggtgacaa aatctgcat	12480
ttggggacat ctgattgtga aagagggagg acagtacact thtagccaca gagactgggg	12540
ctcaccgagc tgaaacctgg tagcactttg gcataacatg tgcatgaccc gtgttcaatg	12600
tctagagatc agtgttgagt aaaacagcct ggtctggggc cgctgctgtc ccacttccc	12660
tcctgtccac cagagggcgg cagagttcct ccaccctgg agcctcccca ggggctgtg	12720
acctccctca gcggggccca cagcccagca gggccaccc tcaccgggt cacctcgcc	12780
cacgtctctc tgcctctcg agctctcac acggactctg tcagctctc cctgcagcct	12840
atcggcggcc cacctgagge ttgtcggcg cccacttgag gctgtcggc tgccctctgc	12900
aggcagctcc tgtccctac acccctcct tccccgggt cagctgaaag ggcgtctccc	12960
agggcagctc cctgtgatct ccaggacagc tcagtctctc acaggtccg acgcccccta	13020
tgctgtcacc tcacagccct gtcattacca ttaactctc agtcccatga agttcactga	13080
gcgcctgtct cccggttaca ggaaaactct gtgacaggga ccacgtctgt cctgtctct	13140
gtggaatccc agggcccagc ccagtgcctg acacggaaca gatgctccat aaatactggt	13200
taaattgtgtg ggagatctct aaaaagaagc atatcacctc cgtgtggccc ccagcagtea	13260
gagtctgttc catgtggaca caggggact ggcaccagca tgggaggagg ccagcaagtg	13320
cccgcggtg cccaggaat gaggcctcaa ccccagagc ttcagaaggg aggacagagg	13380
cctgcaggga atagatctc cggcctgacc ctgcagccta atccagagtt cagggtcagc	13440
tcacaccagc tcgaccctgg tcagatccc tagggcagtt ccagacaagg ccggaggtct	13500
cctcttgccc tccagggggt gacattgcac acagacatca ctcaggaaac ggattccct	13560
ggacaggaac ctggctttgc taaggaagtg gaggtggagc ctggtttcca tcccttgctc	13620

caacagaccc ttctgatctc tccacatac ctgctctgtt cttttctggg tcctatgagg	13680
accctgttct gccaggggtc cctgtgcaac tccagactcc ctctgggtac caccatgggg	13740
aaggtgggggt gatcacagga cagtcagcct cgcagagaca gagaccaccc aggactgtca	13800
gggagaacat ggacaggccc tgagccgcag ctacagccaac agacacggag agggagggtc	13860
cccctggagc cttccccaag gacagcagag ccagaggtca cccacctccc tccaccacag	13920
tcctctcttt ccaggacaca caagacacct ccacctccac atgcaggatc tggggactcc	13980
tgagacctct gggcctgggt ctccatccct gggtcagtgg cggggttggt ggtactggag	14040
acagaggggt ggtccctccc cagccaccac ccagttagcc tttttctagc cccagagcc	14100
acctctgtca ctttctgtt gggcatcatc ccacctccc agagccctgg agagcatggg	14160
gagaccggg accctgctgg gtttctctgt cacaaggaa aataatcccc ctggtgtgac	14220
agaccgaagg acagaacaca gcagaggtca gactgggga agacaggttg tcctcccagg	14280
ggatgggggt ccatecacct tgccgaaaag atttgtctga ggaactgaa atagaaggga	14340
aaaaagagga gggacaaaag aggcagaaat gagaggggag gggacagagg acacctgaat	14400
aaagaccaca cccatgaccc acgtgatgct gagaagtact cctgccctag gaagagactc	14460
transcription start site agggcagagg gaggaaggac agcagaccag acagtcacag cagccttgac aaaacgttcc	14520
tggaactcaa gctcttctcc acagaggagg acagagcaga cagcagagac catggagtct	14580
ccctcggccc ctccccacag atggtgcatc ccctggcaga ggctcctgct cacaggtgaa	14640
gggaggacaa cctgggagag ggtgggagga gggagctggg gtctcctggg taggacaggg	14700
ctgtgagacg gacagagggc tcctgttgga gcctgaatag ggaagaggac atcagagagg	14760
gacaggagtc acaccagaaa aatcaaattg aactggaatt ggaaggggc aggaaaacct	14820
caagagttct attttctag ttaattgtca ctggccacta cgtttttaaa aatcataata	14880
actgcatcag atgacacttt aaataaaaac ataaccaggg catgaaacac tgtcctcatc	14940
cgctaccgc ggacattgga aaataagccc caggctgtgg agggccctgg gaacctcat	15000
gaactcatcc acaggaatct gcagcctgtc ccaggcactg gggtgcaacc aagatc	15056

Mucin-TRE SEQ ID NO:___

5 cgagcggccc ctcagcttcg ggcgccagcc ccgcaaggct cccggtgacc actagagggc 60
gggaggagct cctggccagt ggtggagagt ggcaaggaag gaccctaggg ttcacaggag
120
10 cccagggttta ctcccttaag tggaaatttc ttccccact cctccttggc tttctccaag
180
gaggggaaccc aggctgctgg aaagtccggc tggggcgggg actgtgggtt caggggagaa
240
15 cgggggtgtg aacgggacag ggagcgggta gaagggtggg gctattccgg gaagtgggtg
300
ggggagggag cccaaaacta gcacctagtc cactcattat ccagccctct tatttctcgg
360
20 ccgctctgct tcagtggacc cggggagggc ggggaagtgg agtgggagac ctaggggtg
420
gcttcccgac cttgctgtac aggacctga cctagctggc tttgttcccc atccccacgt
480
tagttgttgc cctgaggcta aaactagagc ccaggggccc caagttccag actgcccctc
540
30 cccctcccc cggagccagg gagtgggttg tgaaaggggg aggccagctg gagaacaaac
600
gggtagtcag ggggttgagc gattagagcc cttgtaccct acccaggaat ggttggggag
660
35 gaggaggaag aggtaggagg taggggaggg ggcgggggtt tgtcacctgt cacctgctcg
720
ctgtgcctag ggcgggcggg cggggagtgg ggggaccggt ataaagcggg aggcgcctgt
780
40 gccgcctcca cctctcaagc agccagcgcc tgccgaatc tggtctgccc cctccccacc
840
45 catttcacca ccaccatg
858

50

α FP-TRE SEQ ID NO:___

gaattcttag aaatatgggg gtaggggtgg tgggtgtaat tctgttttca ccccataggt 60
55 gagataagca ttgggttaaa tgtgctttca cacacacatc acatttcata agaattaagg
120
aacagactat gggctggagg actttgagga tgtctgtctc ataacacttg ggttgatatc
180

gttctatggg gcttgtttta agcttggcaa cttgcaacag ggttcactga ctttctcccc
240

5 aagcccaagg tactgtcctc ttttcatatc tgttttgggg cctctggggc ttgaatatct
300

gagaaaaatat aaacatttca ataatgttct gtggtgagat gagtatgaga gatgtgtcat
360

10 tcatttgtat caatgaatga atgaggacaa ttagtgtata aatccttagt acaacaatct
420

gagggtaggg gtggtactat tcaatttcta tttataaaga tacttatttc tatttattta
480

15 tgcttgtgac aaatgttttg ttcgggacca caggaatcac aaagatgagt ctttgaattt
540

20 aagaagttaa tgggtccagga ataattacat agcttacaaa tgactatgat ataccatcaa
600

acaagagggt ccatgagaaa ataatctgaa aggtttaata agttgtcaaa ggtgagaggg
660

25 ctcttctcta gctagagact aatcagaaat acattcaggg ataattattt gaatagacct
720

taagggttgg gtacattttg ttcaagcatt gatggagaag gagagtgaat atttgaaaac
780

30 attttcaact aaccaaccac ccaatccaac aaacaaaaaa tgaaaagaat ctgagaaaca
840

gtgagataag agaaggaatt ttctcacaac ccacacgtat agctcaactg ctctgaagaa
900

35 gtatatatct aatatttaac actaacatca tgctaataat gataataatt actgtcattt
960

40 tttaatgtct ataagtacca ggcattttaga agatattatt ccatttatat atcaaaataa
1020

acttgagggg atagatcatt ttcattgat atgagaaaaa ttaaaaacag attgaattat
1080

45 ttgcctgtca tacagctaata aattgaccat aagacaatta gatttaaatt agttttgaat
1140

50 ctttctaata ccaaagttca gtttactggt ccatgttgct tctgagtggc ttcacagact
1200

tatgaaaaag taaacggaat cagaattaca tcaatgcaaa agcattgctg tgaactctgt
1260

55 acttaggact aaactttgag caataacaca catagattga ggattgtttg ctgttagcat
1320

60 acaaactctg gttcaaagct cctcttttatt gcttgtcttg gaaaatttgc tgttcttcat
1380

ggtttctctt ttcactgcta tctatcttct tcaaccactc acatggctac aataactgtc
 1440
 5 tgcaagctta tgattcccaa atatctatct ctagcctcaa tcttggtcca gaagataaaa
 1500
 agtagtattc aaatgcacat caacgtctcc acttgaggagg cttaaagacg tttcaacata
 1560
 10 caaaccgggg agttttgctt ggaatgtttc ctaaaatgtg tcctgtagca cataggggtcc
 1620
 tcttggtcct taaaatctaa ttacttttag ccagtgctc atcccaccta tggggagatg
 1680
 15 agagtgaaaa gggagcctga ttaataatta cactaagtca ataggcatag agccaggact
 1740
 20 gtttgggtaa actggtcact ttatcttaaa cttaatatat ccaaaactga acatgtactt
 1800
 agttactaag tctttgactt tatctcattc ataccactca gctttatcca ggccacttat
 1860
 25 ttgacagtat tattgcgaaa acttcctaac tggctcctt atcatagtct tatccccctt
 1920
 tgaaacaaaa gagacagttt caaaatacaa atatgatttt tattagctcc cttttgttgt
 1980
 30 ctataatagt ccgagaagga gttataaact ccatttaaaa agtctttgag atgtggccct
 2040
 tgccaacttt gccaggaatt cccaatatct agtattttct actattaac tttgtgcctc
 2100
 35 ttcaaaactg cattttctct cattccctaa gtgtgcattg ttttccctta ccggttggtt
 2160
 40 tttccaccac cttttacatt ttcttgaac actataccct cctcttcat ttggcccacc
 2220
 tctaattttc tttcagatct ccatgaagat gttacttctt ccaggaagcc ttatctgacc
 2280
 45 cctccaaaga tgtcatgagt tcctcttttc attctactaa tcacagcatc catcacacca
 2340
 50 tgttgtgatt actgatacta ttgtctgttt ctctgattag gcagtaagct caacaagagc
 2400
 tacatgggtc ctgtctcttg ttgtgatta ttcccatcca aaaacagtgc ctggaatgca
 2460
 55 gacttaacat tttattgaat gaataaataa aaccccatct atcgagtgtc actttgtgca
 2520
 60 agaccgggtt ctgaggcatt tatatttatt gatttattta attctcattt aaccatgaag
 2580

gaggctactat cactatcctt attttatagt tgataaagat aaagcccaga gaaatgaatt
2640

5 aactcaccca aagtcattgta gctaagtgc aggcaaaaa ttcaaaccag ttccccaact
2700

ttacgtgatt aatactgtgc tatactgcct ctctgatcat atggcatgga atgcagacat
2760

10 ctgctccgta aggcagaata tggaaggaga ttggaggatg acacaaaacc agcataatat
2820

15 cagaggaaaa gtccaaacag gacctgaact gatagaaaag ttgttactcc tgggtgtagtc
2880

gcacgcacat cttgatgaac tgggtggctga cacaacatac attggcttga tgtgtacata
2940

20 ttattttagt ttgtgtgtgt atttttatat atatatttgt aatattgaaa tagtcataat
3000

ttactaaagg cctaccattt gccaggcatt ttacatttg tccctctaa tcttttgatg
3060

25 agatgatcag attggattac ttggccttga agatgatata tctacatcta tatctatatc
3120

30 tatatctata tctatatcta tatctatatc tatatctata tatgtatatc agaaaagctg
3180

aaatatgttt tgtaaagtta taaagatttc agactttata gaatctggga tttgccaaat
3240

35 gtaacccctt tctctacatt aaacccatgt tggaacaaat acatttatta ttcattcatc
3300

aaatgttgct gagtcctggc tatgaaccag aactgtgaa agcctttggg atattttgcc
3360

40 catgcttggg caagcttata tagtttgctt cataaaactc tatttcagtt cttcataact
3420

45 aatacttcat gactattgct tttcaggtat tccttcataa caaatacttt ggctttcata
3480

tatttgagta aagtcacctt tgaggaagag tagaagaact gcactttgta aatactatcc
3540

50 tggaatccaa acggatagac aaggatggtg ctacctcttt ctggagagta cgtgagcaag
3600

gcctgttttg ttaacatggt ccttaggaga caaaacttag gagagacacg catagcagaa
3660

55 aatggacaaa aactaacaaa tgaatgggaa ttgtacttga ttagcattga agaccttggt
3720

60 tatactatga taaatgtttg tatttgctgg aagtgtact gacggtaaac ctttttgggt
3780

taaatgtgtg ccctagtagc ttgcagtatg atctatTTTT taagtactgt acttagctta
3840

5 tttaaaaatt ttatgtttta aattgcatag tgctctttca ttgaagaagt tttgagagag
3900

agatagaatt aaattcactt atcttaccat ctagagaaac ccaatgttaa aactttgttg
3960

10 tccattatTT ctgtctttta ttcaacattt tttttagagg gtgggaggaa tacagaggag
4020

15 gtacaatgat acacaaatga gagcactctc catgtattgt tttgtcctgt ttttcagtta
4080

acaatatatt atgagcatat ttccatttca ttaaataattc ttccacaaag ttatTTTgat
4140

20 ggctgtatat caccctactt tatgaatgta ccatattaat ttatttcctg gtgtgggtta
4200

tttgatttTTa taatcttacc tttagaataa tgaaacacct gtgaagcttt agaaaatact
4260

25 ggtgcctggg tctcaactcc acagattctg atttaactgg tctgggttac agactaggca
4320

30 ttgggaattc aaaaagttcc cccagtgatt ctaatgtgta gccaagatcg ggaacccttg
4380

tagacaggga tgataggagg tgagccactc ttagcatcca tcatttagta ttaacatcat
4440

35 catcttgagt tgctaagtga atgatgcacc tgaccactt tataaagaca catgtgcaaa
4500

taaaattatt ataggacttg gtttattagg gcttgtgctc taagttttct atgttaagcc
4560

40 atacatcgca tactaaatac tttaaaatgt accttattga catacatatt aagtgaaaag
4620

45 tgtttctgag ctaaacaatg acagcataat tatcaagcaa tgataatttg aaatgaattt
4680

attattctgc aacttaggga caagtcactc ctctgaattt tttgtacttt gagagtattt
4740

50 gttatatTTg caagatgaag agtctgaatt ggtcagacaa tgtcttTgtg gcctggcata
4800

tgataggcat ttaatagttt taaagaatta atgtatttag atgaattgca taccaaattc
4860

55 gctgtctttt ctttatggct tcattaactt aatttgagag aaattaatta ttctgcaact
4920

60 tagggacaag tcatgtcttt gaatattctg tagtttgagg agaataTTtg ttatatTTgc
4980

